

Article

Not peer-reviewed version

Enhancing Parameters Tuning of Overlay Models with Ridge Regression: Addressing Multicollinearity in High-Dimensional Data

[Aris Magklaras](#) , [Christos Gogos](#) ^{*} , [Panayiotis Alefragis](#) , [Alexios Birbas](#)

Posted Date: 13 September 2024

doi: 10.20944/preprints202409.1046.v1

Keywords: Overlay Modeling; Photolithography; Parameters Tuning; EUV lithography; Semiconductor Manufacturing; Yield Results



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Enhancing Parameters Tuning of Overlay Models with Ridge Regression: Addressing Multicollinearity in High-Dimensional Data

Aris Magklaras ¹, Christos Gogos ^{2,*}, Panayiotis Alefragis ³ and Alexis Birbas ¹

¹ Department of Electrical and Computer Engineering, University of Patras, 26504 Patras, Greece; a.magklaras@upnet.gr (A.M.); birbas@ece.upatras.gr (A.B.)

² Department of Informatics and Telecommunications, University of Ioannina, 47100 Arta, Greece; cgogos@uoi.gr

³ Department of Electrical and Computer Engineering, University of Peloponnese, 26334 Patras, Greece; alefrag@uop.gr

* Correspondence: cgogos@uoi.gr; Tel.: +30-6977311944

Abstract: Extreme Ultraviolet (EUV) photolithography process, is a cornerstone of semiconductor manufacturing and operates under demanding precision standards realized via nanometer-level overlay (OVL) error modeling. This procedure allows the machine to anticipate and correct OVL errors before impacting the wafer, thereby facilitating near-optimal image exposure while simultaneously minimizing the overall OVL error. Such models are usually high dimensional and exhibit rigorous statistical phenomena such as collinearities that play a crucial role in the process of tuning their parameters. Ordinary Least Squares (OLS) is the most widely used method for parameters tuning of Overlay models but in most cases it fails to compensate for such phenomena. In this paper we propose the usage of Ridge Regression, a widely known Machine Learning (ML) algorithm especially suitable for datasets that exhibit high multicollinearity. The proposed method has been applied in perturbed data from a 300 mm wafer fab and the results show reduced residuals when ridge regression is applied instead of OLS.

Keywords: overlay modeling; photolithography; parameters tuning; EUV lithography; semiconductor manufacturing; yield results

1. Introduction

According to Moore's law, "the number of transistors that can be placed on a chip doubles every 24 months" [1]. Maintaining Moore's law is highly challenging, because it calls for continuous advancement in a very complex industry. Despite this, recent developments in the production of integrated circuits (ICs), particularly in photolithography, have enabled the industry to keep up with the high process standards. Owing to its role in transferring a desired pattern to a photosensitive material on the wafer surface (photocurable material; most frequently, commercial photo resist; [2]) by exposing it to ultraviolet (UV) or Extreme-UV (EUV) light, photolithography [3] is a crucial component of the entire process. Wafers with a high overlay (OVL) and small critical dimension (CD) are the result of a successful photolithography process. The smaller the exposed pattern and the smaller the exposed ICs on the wafer, the better the OVL and the smaller the CD. Therefore, photolithography is the most crucial step in the production of integrated circuits consisting the base mechanism that supports Moore's law.

In terms of hardware and software, photolithography machines rank among the most complex machines in the market. It is highly challenging to orchestrate and operate the machine in such a way that it satisfies the extremely stringent OVL and CD standards given the machine's more than 50 million lines of code and thousands of hardware modules. This precision is impossible for the hardware alone. Software control techniques that cope with hardware imperfections are essential components of the machine. Software will enable the machine to meet the OVL and CD KPIs, which will confirm the machine's quality.

The fundamental premise underlying this is that the machine can precisely model anticipated wavefront aberrations at the nanoscale level. The software then adjusts the associated machine knobs, such as mirror locations, such that it compensates for the anticipated aberrations because it knows

what to expect. The required pattern is then exposed to the fewest possible flaws because the predicted fingerprint is rectified before it reaches the wafer surface. Under these circumstances, we recognize that one of the most important and difficult responsibilities of the photolithography process is the ability to create precise models. These models are the primary tools for successful exposure.

The process of describing the spatial changes in the overlay (OVL) of the features being printed is known as OVL fingerprint estimation (FE) in photolithography. These discrepancies can be caused by several factors, including flaws in the mask or the lithography procedure itself. The OVL of the features at various positions on the wafer are often measured using specialized metrology instruments that can estimate fingerprints. Information regarding the spatial changes in the OVL is then extracted from the resulting data and utilized to generate a “fingerprint” of the lithography process. Modeling the OVL is an essential step in the FE process. OVL modeling is the process of creating mathematical and statistical models to forecast how various process parameters affect OVL. By modifying the lithography process parameters in real time to meet the necessary OVL criteria, these models can be used to improve the lithography process. To ensure a high yield and reliable performance of the lithography process, FE estimation is a crucial tool.

A polynomial model is a mathematical or statistical representation of a system or process and polynomial. The basis functions and parameters are the two main components of the model. The structure of the problem and the properties of the data being modeled influence the basis functions that are used. On the other hand, model parameters are the estimated values that are used to define the model and are derived from the data. The precise values of the basis functions that best suit the data are determined by the parameters. In our specific use case, we need to provide the model for OVL. A linear model for OVL is defined by:

$$m(x,y) = \Phi \times p$$

Φ is the information matrix which consists of the basis function $\phi(x,y)$ and p are the parameters of our model. The OVL polynomial has P coefficients or parameters. As described above, the information matrix Φ is already known to us and in that case the goal of FE is to estimate the parameters p of our model.

Overlay control is a critical part that enables the exposure system to successfully imprint the complex patterns and meet the strict requirements of modern IC designs. The current photolithographic systems manage to successfully control overlay via a combination of advanced process control (APC) and metrology modules. Metrology contributes into defining and tuning of the overlay models (so that they accurately describe the expected systematic and non-systematic overlay errors [7–9]), as well as relating them to the corresponding controllable knobs of the exposure system [10]. The coupling of overlay error predictions, via overlay models, with machine knobs defines the so called run-to-run (R2R) paradigm of overlay control [11–13]. Obviously, for a R2R process to be successful it is crucial that the overlay models can accurately estimate the expected overlay errors. Extensive research has been conducted in defining overlay models. In [14] multilevel state space models are defined based on existing physics models [15–17] where multilayer, stack up overlay error models are developed. Extensive literature has been conducted for improving overlay modeling process with the focus being either on optimizing the wafer measurements ([18,19]) or in investigating different metrics and cost functions ([20] Zhang *et al.*)

However, it is not sufficient to define the overlay model based only on the underlying physics; it is equally important that the metrology system further tunes the parameters of the overlay model. The Ordinary Least Squares optimization (OLS) method is the most employed regression technique by the current exposure systems. OLS finds the regression coefficients of the overlay model that minimize the residual sum of the squared errors of the difference between the measured and the predicted overlay [21]. In Figure 1, the process of FE is presented. Since the basis functions of the model $\phi(x,y)$ are already predefined, the goal of the FE process is to obtain the best estimation of the model parameters p .

So far, this aspect of the problem, regarding the fitness of the OLS algorithm to model highly dimensional overlay data has not been extensively researched. In this paper we are investigating and proposing the usage of Ridge Regression over the classical OLS algorithm. The remainder of the paper is organized as follows: Section II describes the proposed method and in Section III we show the results of applying the proposed method on an actual industrial process of a 300mm wafers. Finally, Section IV presents the conclusions and potential future work.

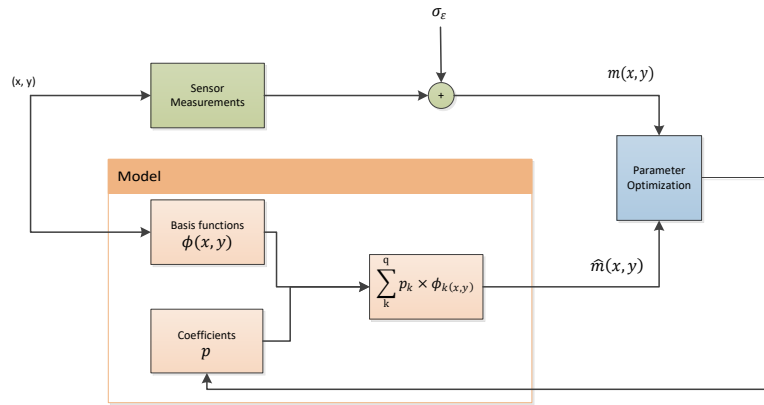


Figure 1. Fingerprint Estimation as a block diagram.

2. Materials & Methods

In this paper we are proposing to use Ridge Regression instead of the classic OLS method for tuning the parameters p of the overlay models. The overlay models as mentioned above consist of a set of basis functions $\phi(x, y)$ that are already predefined based on the physics and specific parameters of the process. Ridge regression or Tikhonov regularization [22] is a statistical method for estimating the parameters of multiple regression models in scenarios where the independent variables are highly correlated.

Ridge regression is particularly useful in ill-posed problems which exhibit multicollinearity in their independent variables. In our case we try to find the vector p of parameters such that:

$$m = \Phi \times p$$

As mentioned before the standard approach is to use the OLS method. OLS seeks to minimize the the sum of the squared residuals:

$$\|\Phi p - m\|_2^2$$

However if no p satisfies the equation or more p do, then the problem is ill-posed and OLS might lead to over- or under-determined system of equations.

Ridge regression adds a regularization term $\|\Gamma x\|_2^2$ for some suitable chosen matrix $\Gamma = \alpha I$. This is known as L_2 regularization. This ensures smoothness and improves the conditioning of the problem. The minimization problem to be solved then is:

$$\|\Phi p - m\|_2^2 + \|\Gamma x\|_2^2$$

And the corresponding parameter vector p :

$$p = (\Phi^T \Phi + \Gamma^T \Gamma)^{-1} \Phi^T m$$

In our proposal we are using $\alpha = 1.0$ which means that $\Gamma = I$.

In Figure 2 the Corelation Matrix of the Overlay X and Overlay Y is presented. There seems to be high correlation between approximately 50 of the 161 parameters. This is an indication (a strong one though) that there is significant collinearity in the parameters and this needs to be further investigated.

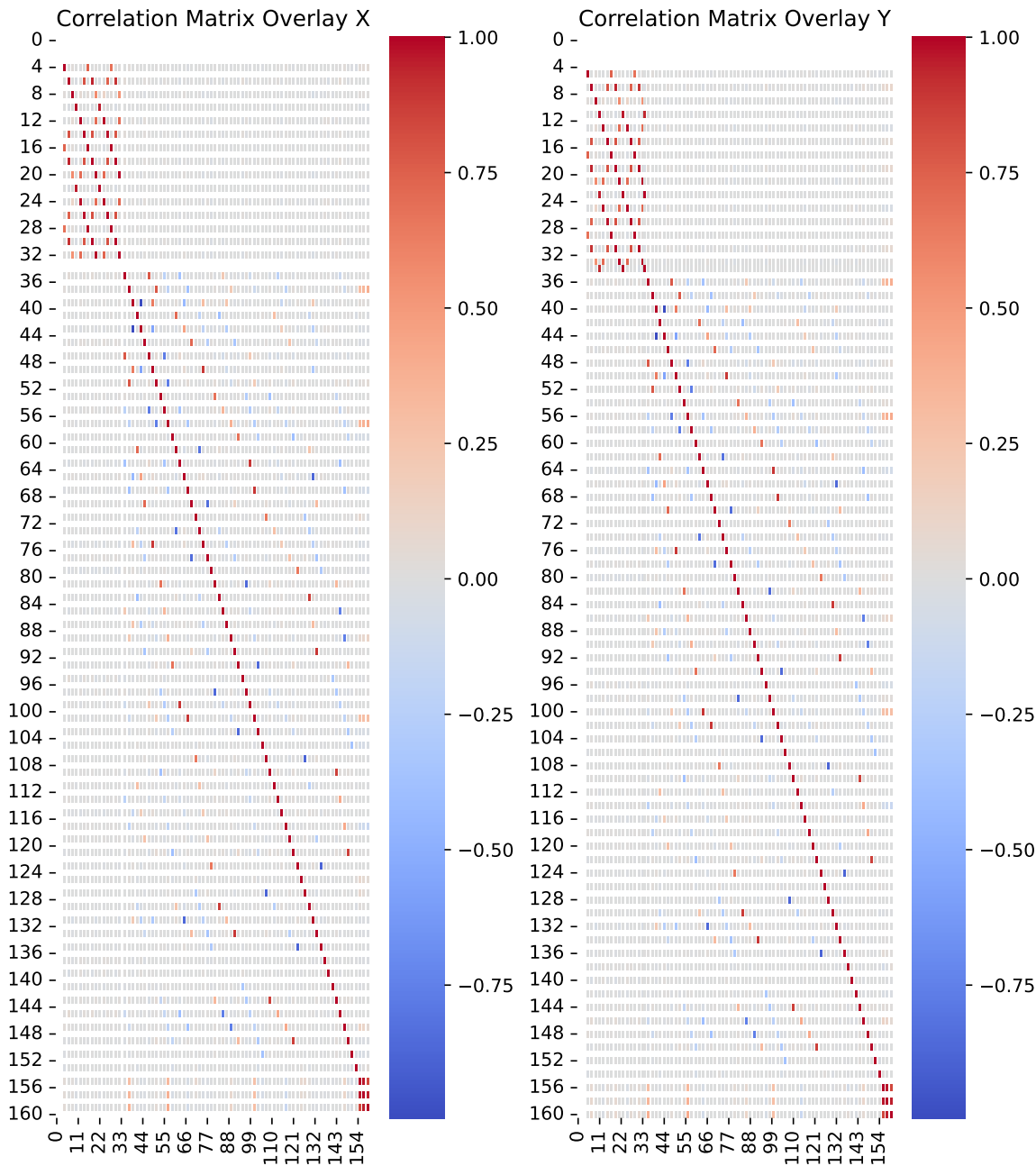


Figure 2. Correlation Matrices for Overlay X and Overlay Y.

The Variance Inflation Factor (VIF) is a statistical measure that quantifies the degree of multicollinearity for each independent variable. VIF is calculated as:

$$VIF = 1/(1 - R^2)$$

with R^2 being the square of the OLS solutions. A value of $VIF > 5$ indicates moderate multicollinearity while $VIF > 10$ indicated high multicollinearity. Performing the VIF check in our OLS solutions for Overlay X and Y results in high multicollinearity in 62/161 parameters for Overlay X and in 109/161

parameters for Overlay Y. So, the conclusion is that multicollinearity needs to be addressed. In this case, indeed the Ridge Regression method should be able to address the issue and improve the parameter estimation method.

3. Results & Discussion

In our experiments we compared the overlay residuals (in X and in Y) for both Ridge Regression and OLS methods. The expected overlay, per field point, is compared to the actual overlay and next a statistical analysis of the results is performed. In assessing the performance we are using the 99.7 percentile and *max* residual metrics. Utilizing these metrics enables a comprehensive understanding of the overlay performance, providing insights into the extent of variability and the upper bounds of error dispersion.

The goal of our method is to accurately model the measured overlay and reproduce it with minimum error. In Figure 3 we can see the measured Overlay X on the left and the modeled Overlay X on the right, using the Ridge Regression Method. In this visual representation we can see that the modeled Overlay is able to capture successfully the expected overlay. Also in areas of the wafer where there seems to be abnormal behavior as on the edges of the wafer, the patterns seems to really match. Similarly in Figure 4, Ridge Regression also performs well on the Overlay Y.

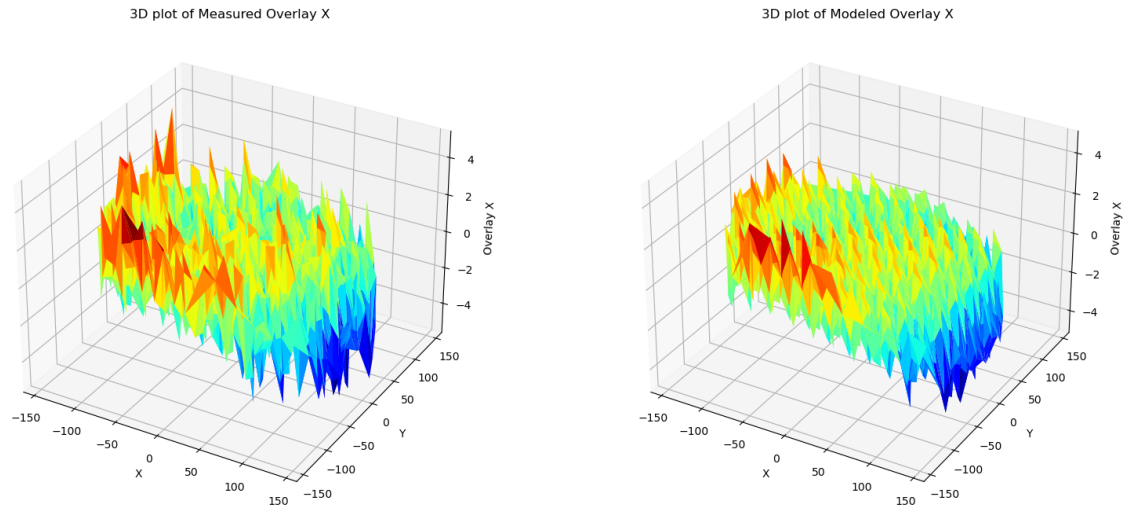


Figure 3. Overlay X - Measured vs Modeled for layer 4.

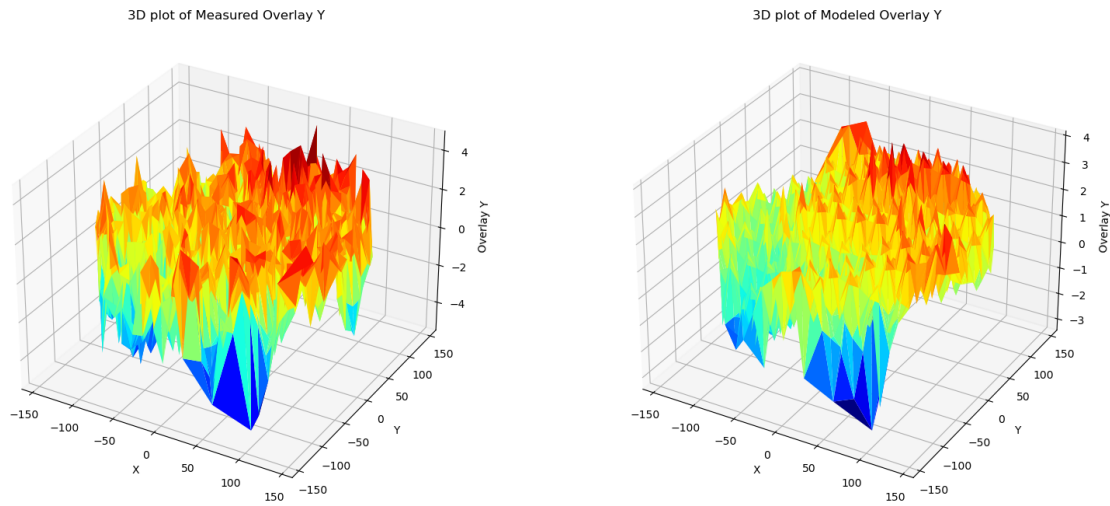


Figure 4. Overlay Y - Measured vs Modeled for layer 4.

The residuals vs. the predicted values for Overlay X and Y are presented in Figure 5 too.

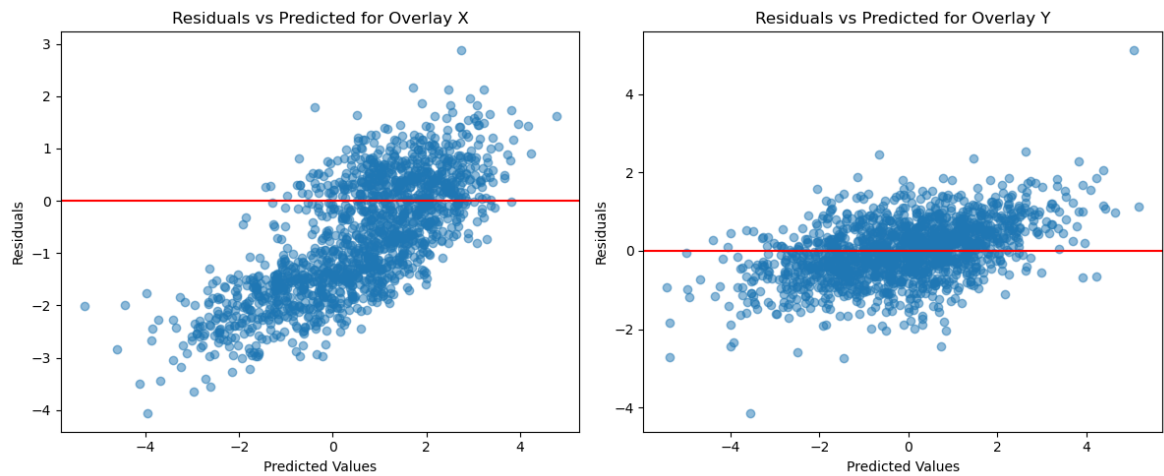


Figure 5. Residuals vs. Predicted for layer 4.

In Figure 6 the Overlay X residuals in 99.7 are presented, for both OLS and Ridge Regression methods. For 8/12 layers ridge regression outperforms OLS method. OLS still is better than Ridge Regression in layers 3, 5, 6 but given the notably high residuals in Ridge Regression we can consider these results as outliers.

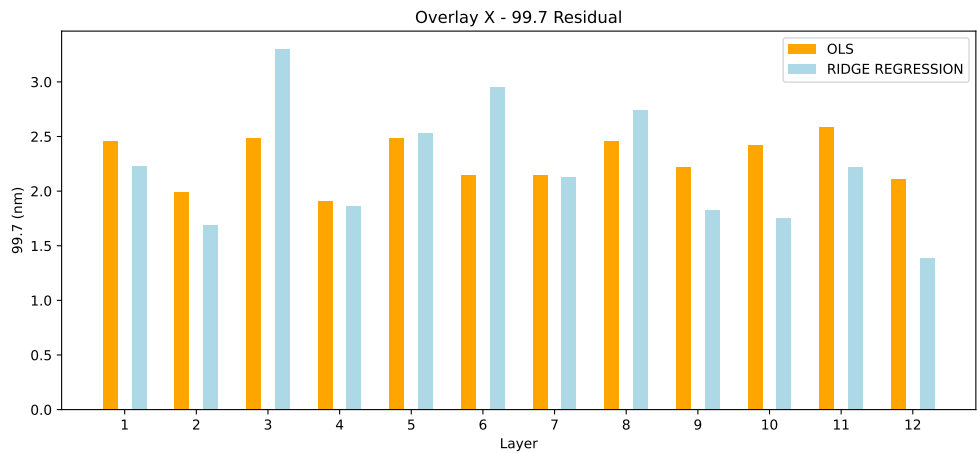


Figure 6. Overlay X - 99.7 Residual.

The max residuals in Overlay X are compared in Figure 7. Here, we observe a different pattern than in 99.7. OLS outperforms the ridge regression in 9/12 layers. From these results we can not yet draw a safe conclusion. It depends which metric we value most (99.7 vs. max) and this actually depends on the use case. However, the superiority of ridge regression in 99.7 only is an interesting conclusion already.

When checking the Overlay in Y we draw a more clear picture of the situation. As presented in Figure 8 Ridge Regression outperforms significantly the OLS in all 12 layers. On average this is a 1.04 nm improvement. In this case the superiority of Ridge Regression is clear.

Similarly, in Overlay MAX we observe Ridge Regression outperforming OLS in all wafers with an average improvement of 1.10 nm.

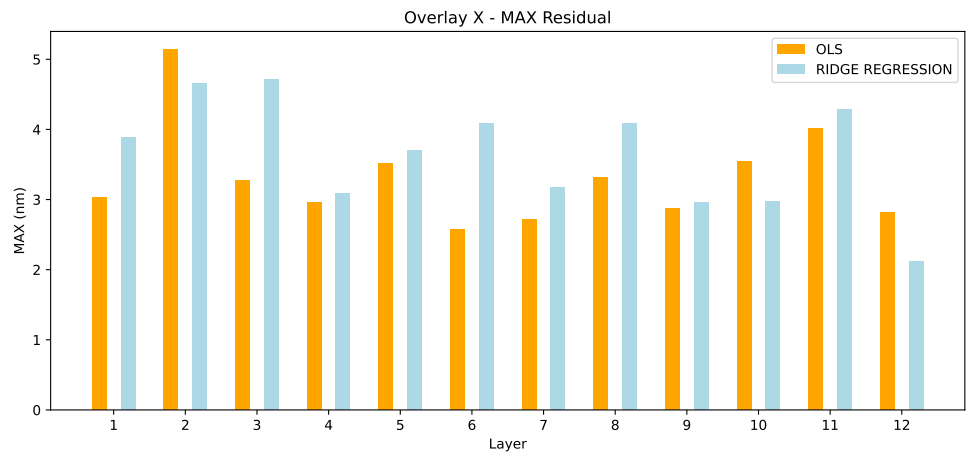


Figure 7. Overlay X - MAX Residual.

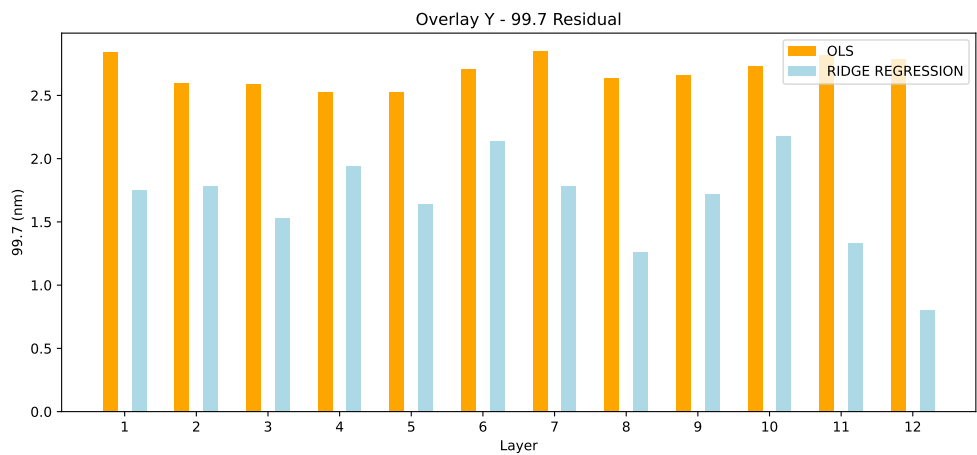


Figure 8. Overlay Y - 99.7 Residual.

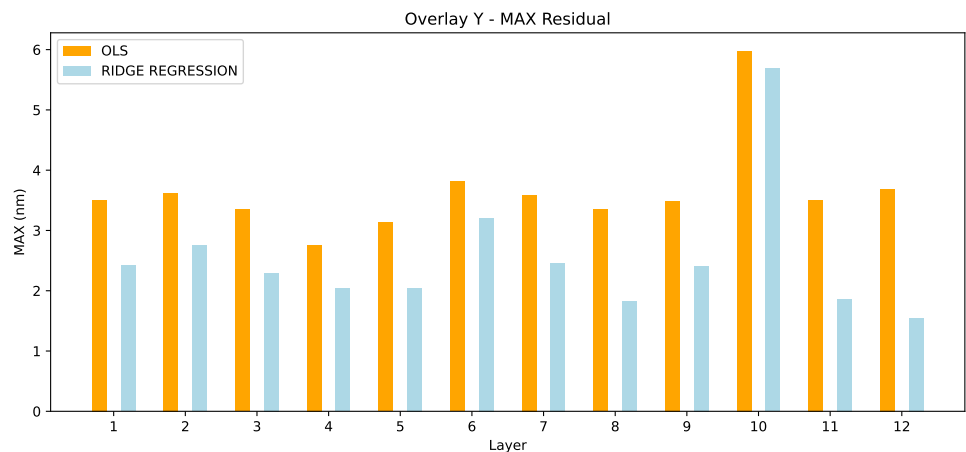


Figure 9. Overlay Y - MAX Residual.

The interesting conclusion, which verifies our hypothesis, is that Ridge Regression significantly outperforms OLS especially in Overlay Y. This makes sense since we already showed that the parameters which exhibit multicollinearity are almost double in Overlay Y than in Overlay X.

In Figures 10 and 11 we present the measured vs. modeled Overlay X and Y for all the 12 layers.

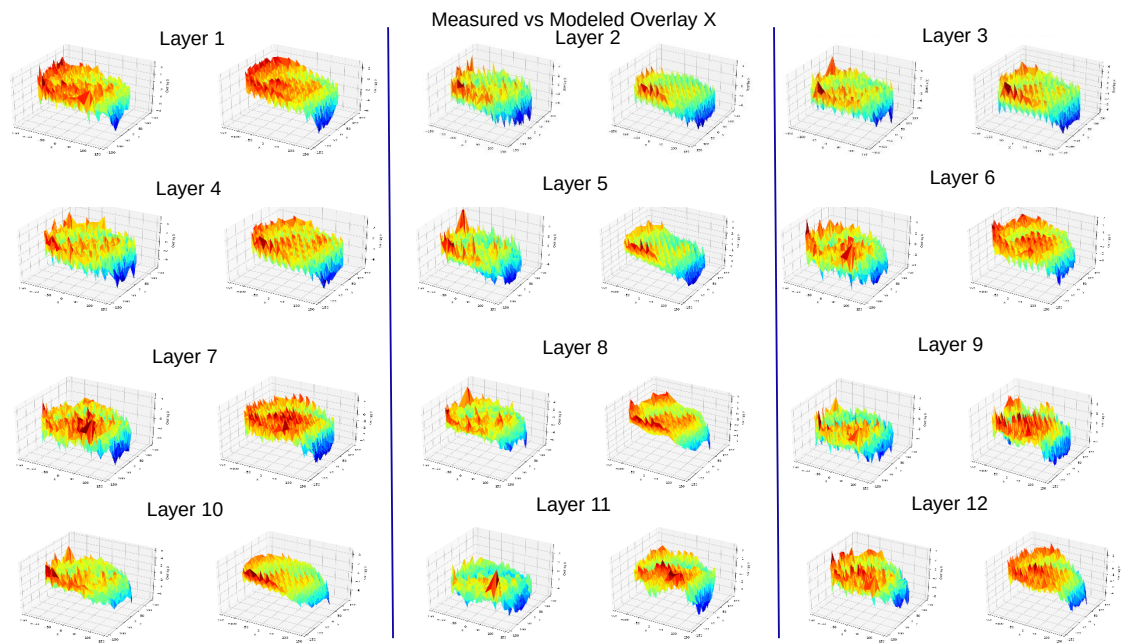


Figure 10. Measured vs. Modeled Overlay X for all 12 layers.

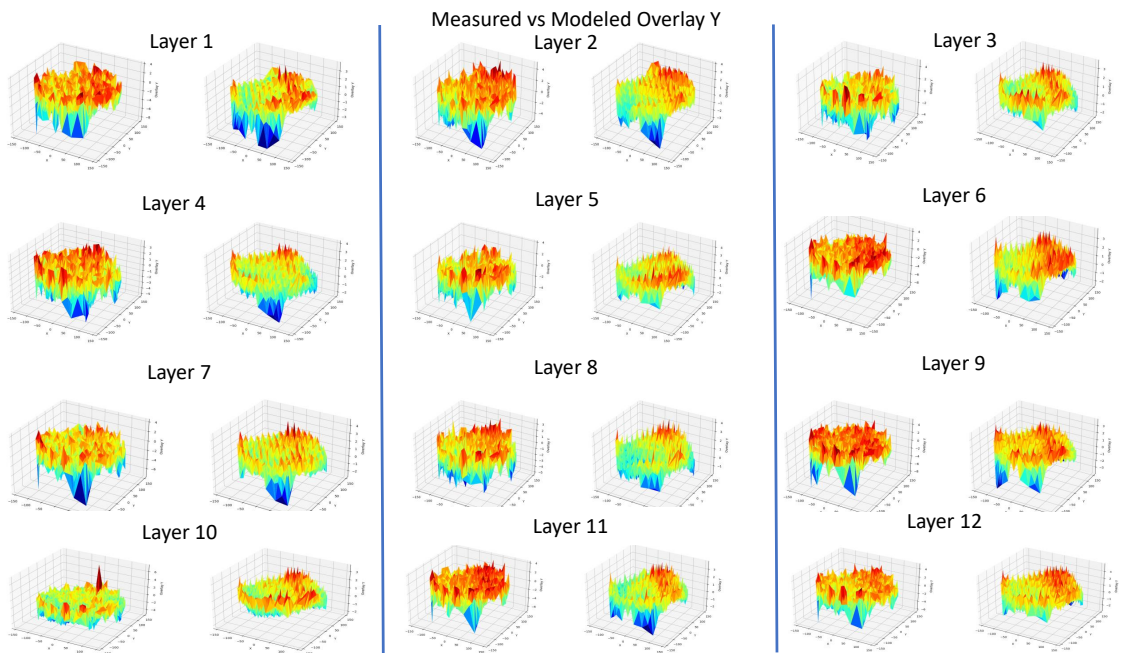


Figure 11. Measured vs. Modeled Overlay Y for all 12 layers.

4. Conclusion & Future Work

Ridge Regression results indicate an obvious superiority over OLS when the data exhibit multicollinearity. Especially in Overlay X where this phenomenon is more intense Ridge Regression outperforms OLS in every single layer. In the end, the overall improvement in 99.7 is 1.04 nm and in the max 1.10 nm. Such performance improvement in sub-nanometer level of precision can not be ignored. It is clear that OLS needs to be enhanced for similar datasets. One potential direction for future work would be the investigation of improving the model itself, either by reducing the features or by dynamically, during wafer production, select the most informative ones depending on the state of the process.

Author Contributions: Conceptualization, A.M.; methodology, A.M., C.G. and P.A.; software, A.M.; writing—original draft preparation, A.M.; writing—review and editing, A.M., P.A. and C.G.; supervision, A.B.; All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data presented in this study are not publicly available due to IP protection reasons.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. "Moore's Law Inspires Intel Innovation", <https://www.intel.com/content/www/us/en/history/museum-gordonmoore-law.html>.
2. Jacobs, I. S. "Fine particles, thin films and exchange anisotropy." *Magnetism* (1963): 271-350.
3. Lee, K. Y., et al. "Micromachining applications of a high resolution ultrathick photoresist." *Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures Processing, Measurement, and Phenomena* 13.6 (1995): 3012-3016.
4. C. A. Mack, *Fundamental Principles Of Optical Lithography: The Science Of Microfabrication*. Chichester, UK: Wiley, 2012.
5. H. Geng, *Semiconductor Manufacturing Handbook*, 2nd ed. New York, NY, USA: McGraw-Hill Educ, 2018.
6. D. Djurdjanovic, L. Mears, F. A. Niaki, A. U. Haq, and L. Li, "State of the art review on process, system, and operations control in modern manufacturing," *ASME J. Manuf. Sci. Eng.*, vol. 140, no. 6, 2018, Art. no. 061010. DOI: 10.1115/1.4038074
7. B. Y. Hsueh, et al., "High order correction and sampling strategy for 45nm immersion lithography overlay control," in *Proc. SPIE Metrol. Inspection Process Control Microlithography XXII*, vol. 6922, Mar. 2008, Art. no. 69222Q. DOI: 10.1117/12.772107
8. D. Choi, et al., "Optimization of high order control including overlay, alignment, and sampling," in *Proc. SPIE*, vol. 6922. San Jose, CA, USA, Feb. 2008, Art. no. 69220P. DOI: 10.1117/12.772274
9. H. Lin, B. Lin, J. Wu, S. Chiu, C.-C. Huang, J. Manka, et al., (2008). Improve overlay control and scanner utilization through high order corrections. *Proceedings of SPIE - The International Society for Optical Engineering*. 6922. DOI: 10.1117/12.772118.
10. M. A. van den Brink, C. G. M. de Mol, and R. A. George, "Matching performance for multiple wafer steppers using an advanced metrology procedure," in *Proc. Integr. Circuit Metrol. Inspection Process Control II*, vol. 0921, Online (Bergh.) p. 180, Jan. 1988., doi:DOI: 10.1117/12.968366
11. A. U. Haq and D. Djurdjanovic, "Robust control of overlay errors in photolithography processes," *IEEE Trans. Semicond. Manuf.*, vol. 32, no. 3, pp. 320–333, Aug. 2019. DOI: 10.1109/TSM.2019.2913391
12. C.-F. Chien, Y.-J. Chen, C.-Y. Hsu, and H.-K. Wang, "Overlay error compensation using advanced process control with dynamically adjusted proportional-integral R2R controller," *IEEE Trans. Autom. Sci. Eng.*, vol. 11, no. 2, pp. 473–484, Apr. 2014. DOI: 10.1109/TASE.2013.2280618
13. F. Tan, T. Pan, Z. Li, and S. Chen, "Survey on run-to-run control algorithms in high-mix semiconductor manufacturing processes," *IEEE Trans. Industr. Inform.*, vol. 11, no. 6, pp. 1435–1444, Dec. 2015. DOI: 10.1109/TII.2015.2490039
14. J. Yu and S. Qin, "Variance component analysis based fault diagnosis of multi-layer overlay lithography processes," *IIE Trans.*, vol. 41, no. 9, pp. 764–775, 2009. DOI: 10.1080/07408170902789076
15. F. He and Z. Zhang, "An empirical study-based state space model for multilayer overlay errors in the step-scan lithography process," *RSC Adv.*, vol. 5, no. 126, 103901–103906, 2015. DOI: 10.1039/C5RA07164J
16. F. He and Z. Zhang, "State space model and numerical simulation of overlay error for multilayer overlay lithography processes," in *Proc. 2nd Int. Conf. Image Vis. Comput. (ICIVC)*, 2017, pp. 1123–1127. DOI: 10.1109/ICIVC.2017.7984728
17. M. Khakifirooz, M. Fathi, and C.-F. Chien, "Partially observable Markov decision process for monitoring multilayer wafer fabrication," *IEEE Trans. Autom. Sci. Eng.*, vol. 18, no. 4, pp. 1742–1753, Oct. 2021. DOI: 10.1109/TASE.2020.3017481

18. A. Magklaras, P. Alefragis, C. Gogos, C. Valouxis, and A. Birbas, "A Genetic Algorithm-Enhanced Sensor Marks Selection Algorithm for Wavefront Aberration Modeling in Extreme-UV (EUV) Photolithography," *Information (Basel)*, vol. 14, no. 8, p. 428, 2023., doi:DOI: 10.3390/info14080428
19. A. Magklaras, C. Gogos, P. Alefragis, C. Valouxis, and A. Birbas, "Sampling Points Selection Algorithm For Advanced Photolithography Process," 2022 7th South-East Europe Design Automation, Computer Engineering, Computer Networks and Social Media Conference (SEEDA-CECNSM), Ioannina, Greece, 2022, pp. 1-5, doi: DOI: 10.1109/SEEDA-CECNSM57760.2022.9932920.
20. H. Zhang, T. Feng, and D. Djurdjanovic, "Dynamic Down Selection of Measurement Markers for Optimized Robust Control of Overlay Errors in Photolithography Processes," *IEEE Trans. Semicond. Manuf.*, vol. 35, no. 2, pp. 241–255, 2022., doi:DOI: 10.1109/TSM.2022.3143039
21. S.-C. Horng and S.-Y. Wu, "Compensating the overlay modeling errors in lithography process of wafer stepper," in *Proc. IEEE Conf. Ind. Electron. Appl. (ICIEA)*, Taichung, Taiwan, Jun. 2010, pp. 1399–1404.
22. G.C. McDonald, "Ridge regression.", *Wiley Interdisciplinary Reviews: Computational Statistics* 1.1 (2009): 93-100.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.